



# FG-iRODS

Pierre Gay

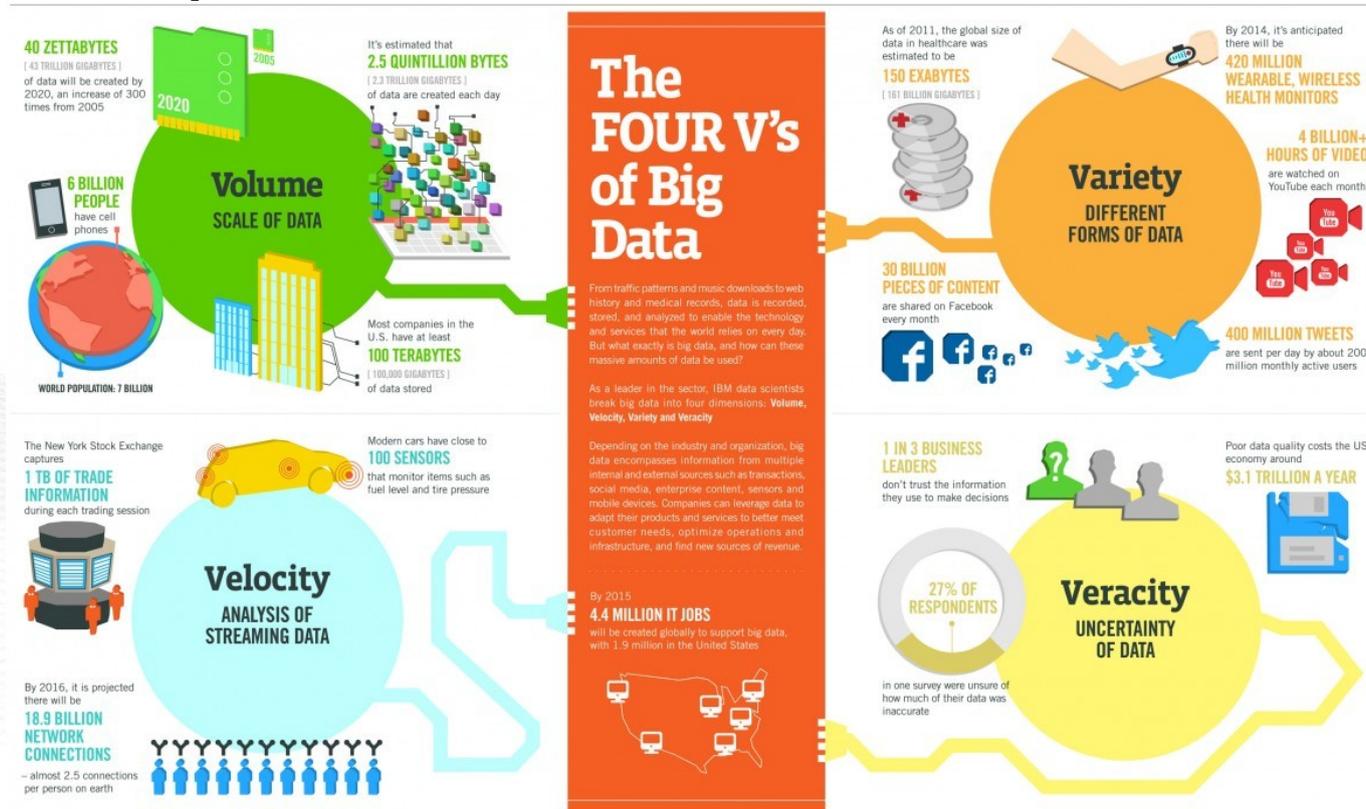
Workshop Opérations France-Grilles 17 juin 2015



# Introduction

# Big Data

- Par définition, le « Big Data » concerne des données particulières





## Fichiers

- Stocker les données scientifiques sous forme de fichiers reste la norme (notamment pour des données stables et de confiance)
- Certaines collections sont en effet volumineuses et nécessitent des traitements importants
- Dropbox ne semble pas adapté à tous les besoins ;)
- Les systèmes de grille comme LCG/EGI sont des succès, mais pas forcément adéquat pour tous les besoins (données à caractère patrimonial, long tail of science)



# iRODS

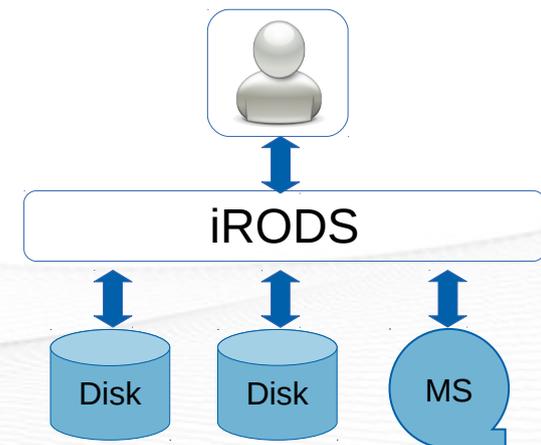
## iRODS

- Un logiciel pour construire sa propre « grille de données »
  - Un catalogue central (LFNs, ACLs, métadonnées)
  - Des ressources de stockage éventuellement distribuées géographiquement (réplication possible)
  - Un « moteur de règles » pour implémenter les politiques de gestion
  - Des logiciels clients accessibles au commun des mortels

## IRODS : iRule Oriented Data System

- Projet démarré en 2006 (basé sur SRB)
- Distribué en Open Source (BSD)
- Développé par le groupe DICE et différents collaborateurs
- <http://irods.org>

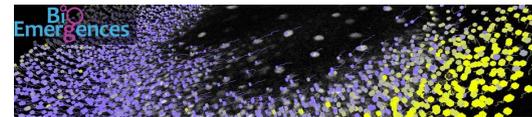
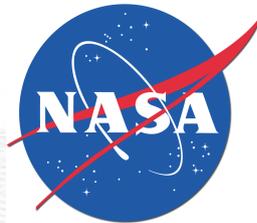
# iRODS



## Fiche technique

- Utilisable depuis un portable jusqu'à de dépôts institutionnels de projets internationaux
- Milliers d'utilisateurs
- Milliards de fichiers et plusieurs péta-octets
- Documentation extensive
  - <https://wiki.irods.org/index.php/Documentation>
- Paquets binaires disponibles
- Version 4 est une recombinaison de 2 branches du logiciel

**{ BnF** Bibliothèque nationale de France

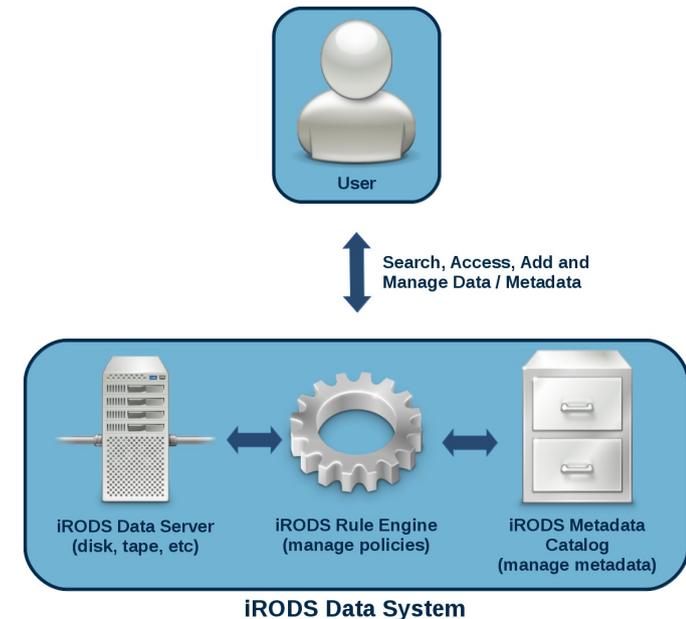


Observatoire astronomique  
de Strasbourg

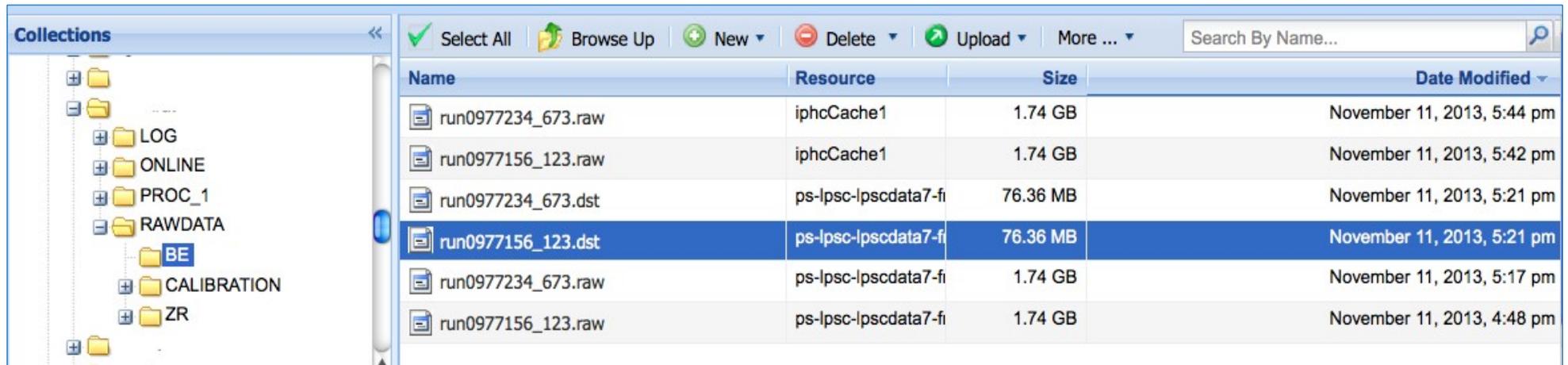


## Sous le capot

- Une zone iRODS comprend :
  - Un catalogue / Base de données
  - Un Moteur de Règles
  - Des Ressources de stockage
- Les serveurs peuvent être distribués géographiquement
- Possibilité d'interconnecter plusieurs zones
- Interfaces utilisateurs : GUI, CLIs, APIs (C, Java, PHP, Python, ...)
- **Transferts de données non chiffrés avant la version 4**



# Interfaces utilisateurs



Name	Resource	Size	Date Modified
run0977234_673.raw	iphcCache1	1.74 GB	November 11, 2013, 5:44 pm
run0977156_123.raw	iphcCache1	1.74 GB	November 11, 2013, 5:42 pm
run0977234_673.dst	ps-lpsc-lpscddata7-fi	76.36 MB	November 11, 2013, 5:21 pm
run0977156_123.dst	ps-lpsc-lpscddata7-fi	76.36 MB	November 11, 2013, 5:21 pm
run0977234_673.raw	ps-lpsc-lpscddata7-fi	1.74 GB	November 11, 2013, 5:17 pm
run0977156_123.raw	ps-lpsc-lpscddata7-fi	1.74 GB	November 11, 2013, 4:48 pm

```
[user ~]$ ils
/frgrid/home/UNECOLLAB/RAWDATA:
C- /frgrid/home/UNECOLLAB/RAWDATA/CALIBRATION
C- /frgrid/home/UNECOLLAB/RAWDATA/BE
C- /frgrid/home/UNECOLLAB/RAWDATA/ZR
[user ~]$ ils -l BE/
/frgrid/home/UNECOLLAB/RAWDATA/BE:
owner  0 ps-lpsc-lpscddata7-fr  80072192 2013-11-11.16:21 & run0977156_123.dst
owner  0 ps-lpsc-lpscddata7-fr 1748189011 2013-11-11.15:48 & run0977156_123.raw
owner  1 iphcCache1          1748189011 2013-11-11.16:42 & run0977156_123.raw
owner  0 ps-lpsc-lpscddata7-fr  80072192 2013-11-11.16:21 & run0977234_673.dst
...
```

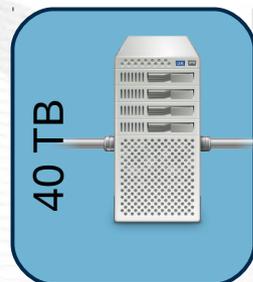


# FG-IRODS

## Infrastructure fédérée FG-iRODS



- Coordonnée par France Grilles
- Fédération des ressources et des efforts
- Pour des utilisateurs de tous les domaines scientifiques
- Destinée aux petits ou moyens projets
- Ouvert aux nouveaux fournisseurs de ressources
- Formation et support aux utilisateurs

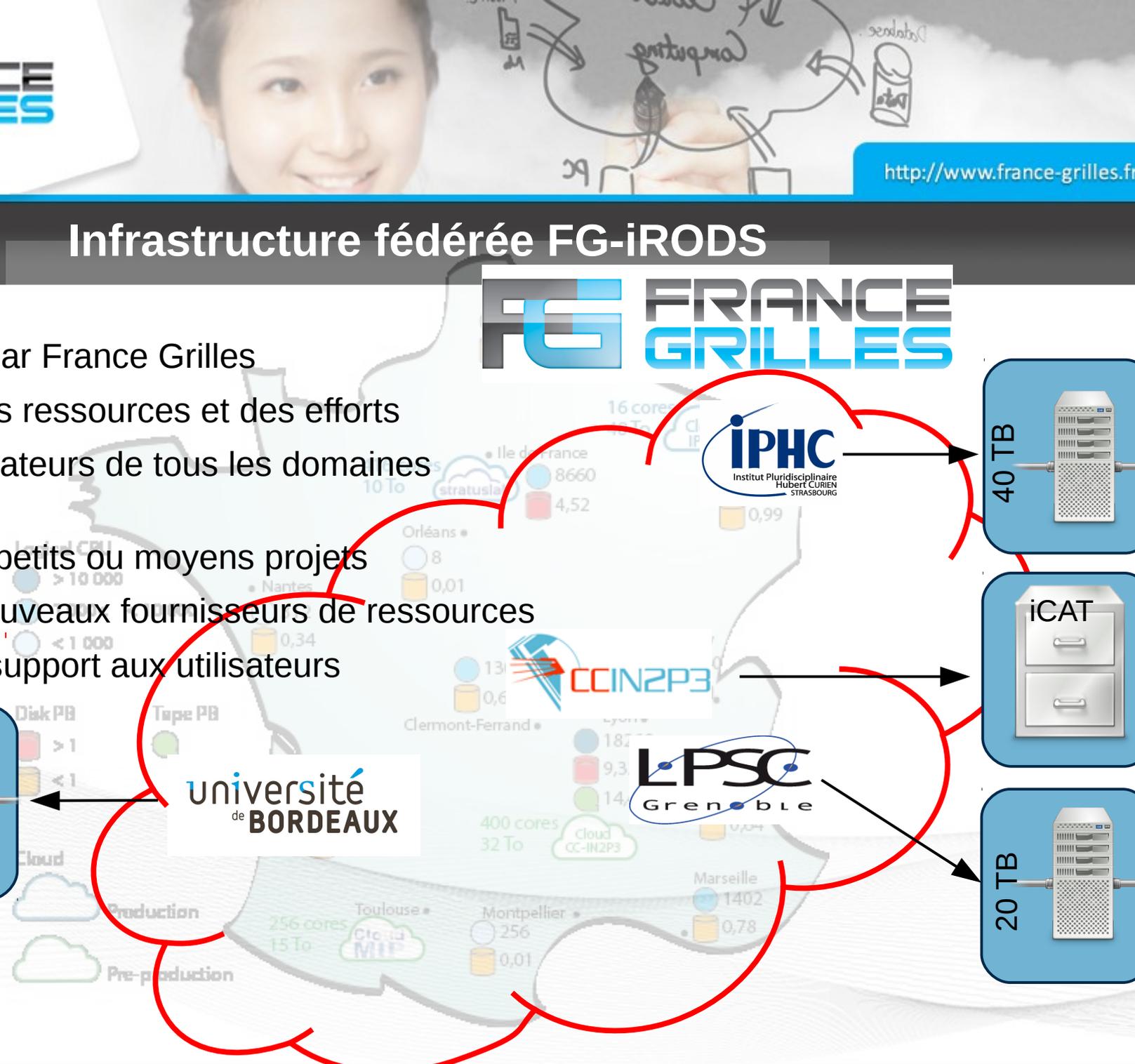
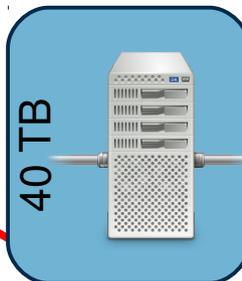


université  
de **BORDEAUX**

**IPHC**  
Institut Pluridisciplinaire  
Hubert CURIE  
STRASBOURG

**CCIN2P3**

**LPSC**  
Grenoble







## L'équipe FG-iRODS

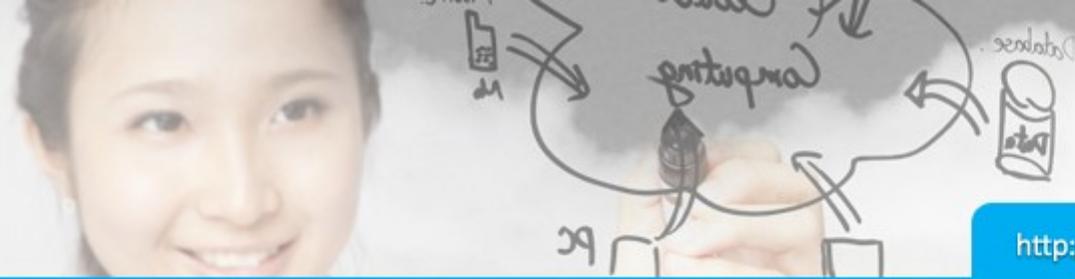
- Yonny CARDENAS (CC-IN2P3, Lyon)
- Jean-Yves NIEF (CC-IN2P3, Lyon)
- Geneviève ROMIER (France Grilles, Lyon)
- Emmanuel MEDERNACH (IPHC, Strasbourg)
- Jérôme PANSANEL (IPHC, Strasbourg)
- Catherine BISCARAT (LSPC, Grenoble)
- David BENABEN (CBIB & INRA, Bordeaux)
- Pierre GAY (MCIA, Bordeaux)

## Les services

- Accueil des utilisateurs
  - Premier contact : [info@france-grilles.fr](mailto:info@france-grilles.fr) et formulaire
  - Pour avoir un compte :
    - inscription dans la VO [vo.france-grilles.fr](http://vo.france-grilles.fr)
    - Signature de l'AUP iRODS : [http://www.france-grilles.fr/IMG/pdf/iRODS\\_Service\\_Policy.pdf](http://www.france-grilles.fr/IMG/pdf/iRODS_Service_Policy.pdf)
  - Mailinglist utilisateurs : [irods-user-l@FRANCE-GRILLES.FR](mailto:irods-user-l@FRANCE-GRILLES.FR)
  - Authentification par certificat et mot de passe
  - Cycle de vie des comptes (revue au bout d'un an)
- Support
  - Documentation : [https://forge.in2p3.fr/projects/france-grilles-documentation/wiki/Utiliser\\_le\\_service\\_iRODS](https://forge.in2p3.fr/projects/france-grilles-documentation/wiki/Utiliser_le_service_iRODS)
  - Forums : <https://forge.in2p3.fr/projects/france-grilles-documentation/boards/101>
  - Mailinglist : [irods-support-l@FRANCE-GRILLES.FR](mailto:irods-support-l@FRANCE-GRILLES.FR)
- Service accessible depuis la grille
- Fonctionnalités à étudier sur demande :
  - Réplication automatique à l'ingestion
  - Checksum automatique à l'ingestion
  - Métadonnées automatiques à l'ingestion
  - Péremption des données

## Réalisations

- Fédération de 100 To de ressources de stockage (iRODS 3.3.1)
- 17 utilisateurs actifs, 2,7 To stockés
- Réelle synergie entre les administrateurs
- AUP publiée
- Plusieurs formations utilisateurs (récemment : Bordeaux 01/2015)
- Clients iRODS installés sur les sites supportant la VO [vo.france-grilles.fr](http://vo.france-grilles.fr)
- Paquets clients iRODS disponibles avec support GSI (deb, rpm) et script d'installation, VM *appliance* disponible
- Interface web disponible
- Documentation utilisateur :  
[https://forge.in2p3.fr/projects/france-grilles-documentation/wiki/Utiliser\\_le\\_service\\_iRODS](https://forge.in2p3.fr/projects/france-grilles-documentation/wiki/Utiliser_le_service_iRODS)



## Perspectives

- Formations à venir : deuxième semestre 2015
- Tests en cours pour éprouver l'infrastructure : E. Medernach
- Comment pérenniser le projet ?
- Approfondir la co-administration des sites
- Attirer d'autres sites dans la collaboration
  - Au moins un projet intéressé
- Développements à étudier :
  - Service de changement de mot de passe iRODS à partir du certificat grille
  - Cycle de vie des données
    - Destruction des données après fermeture de compte utilisateur
    - Nettoyage du trash
    - Vérification périodique de l'intégrité des données



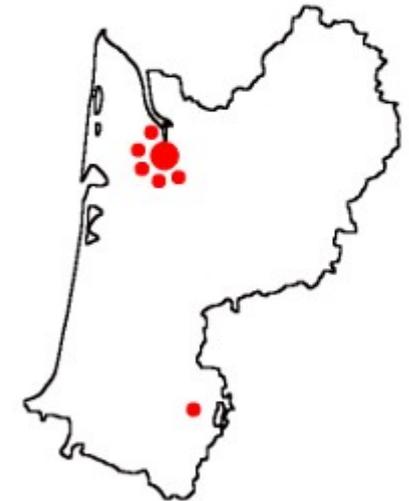
# Un projet lié : MCIA-IRODS

## Mon iRODS à moi

- Projet MCIA-iRODS
  - Débuté en 2013
  - Mésocentre + 6 partenaires
  - 1+ Po de capacité disque
  - Cible : 300+ utilisateurs
- Encore en bêta test :(
- Pas mal d'études au cours du projet
  - sur la co-administration
  - sur les fonctionnalités pour les utilisateurs
- Certains partenaires contribuent des ressources sur financement propre dans l'infrastructure



# iRODS





# Interactions possibles MCIA-iRODS / FG-iRODS

## • Utilitaires :

- <https://github.com/mesocentre-mcia/mcia-irods-client-install>
  - Script d'installation des clients + utilitaires
- <https://github.com/mesocentre-mcia/mcia-irods-utils>
  - Commandes : ilsw, irmw, igetw, idu, irepl-check, imeta-coll, iextract

## • Fonctionnalités :

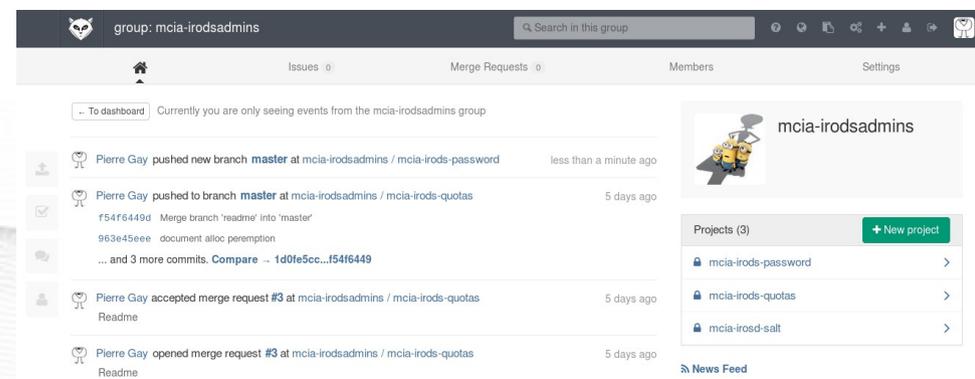
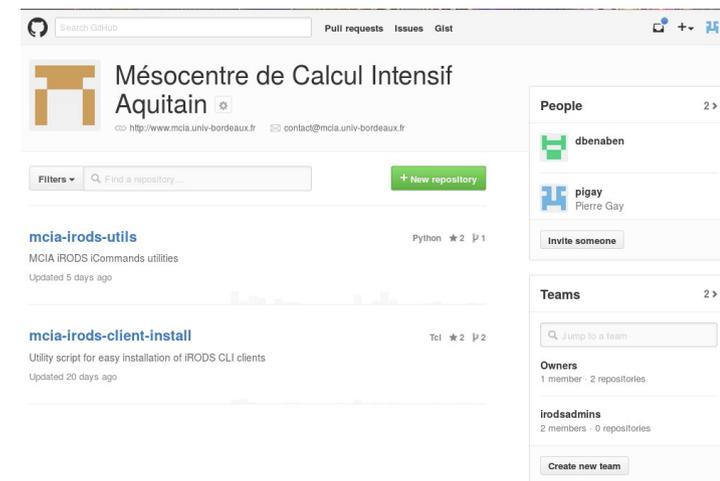
- Dispositif de réplication automatique configurable
- Checksum à l'ingestion
- Nettoyage du trash utilisateur
- Gestion des quotas (quotas de projets)

## • Outils d'administration

- Co-administration : saltstack
- Procédures de contrôle d'intégrité des répliques
- Service de changement de mot de passe iRODS

## • Prospective

- Fédération des zones MCIA-iRODS / FG-iRODS
- Essais iRODS 4 ?





## Questions

- Merci de votre attention